FS-DS based Multi-sensor Data Fusion

Shunyuan Sun, Jing Gao, Minfang Chen, Baoguo Xu*
School of IOT Engineering,
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education).
Jiangnan University.
Wuxi, 214122, P.R. China.
E-mail: 54robin@163.com

Zhiguo Ding
School of Electrical, Electronic, and Computer Engineering.
Newcastle University, UK
E-mail: yantaissy@gmail.com

***Abstract*—considering the problem of the uncertainty of data in information gathering system, a multi-sensor information fusion method based on fuzzy set and evidence theory (FS-DS) is proposed. The fuzzy support probability of the uncertain information is defined by making using of the correlation function, Then it is received the credibility of the information measured by each sensor form the membership function. And , will the support and confidence into basic probability function. Finally, the sensors with higher measurement precision are identified by D-S evidence combination. The proposed method can improve the problem of the basic probability assignment function is difficult to be determined and the calculation of degree of mutual support is absolute are improved. The practical application verifies that the fusion result has higher accuracy and reliability.**

***Index Terms*—multi-sensor, data fusion, evidence theory, fuzzy set**

I. INTRODUCTION

Multi-sensor data fusion[1] can comprehensive utilize the data measured by each sensor overcome the uncertainty of measured data, compared with a single sensor, it improves the effective performance of entire sensor system, and can describe the tested object [2] more accurately. And in the multi-sensor information gathering system, it is inevitable affected by many factors, such as the sensor precision, transmission error, the environmental noise, and human disturbance, those factors will interest uncertainty [3] of the measured data. Because of the persons of utilizing information can't accurate judge the true state of things which reflected from the information, so lead to the deviation of subjective and itself. At present, how to comprehensive process the uncertainty of the information and accurate judge the true state of things using the advanced data fusion method, have received wide attention of related researchers.

In view of the above questions, a multi-sensor information fusion[4] method based on fuzzy set and evidence theory (FS-DS) is proposed in this paper. This method express the mutual support degree of the measured data with the size of the fuzzy believe distance measurement, abandon the absolution of support probability selection in the past[5], define fuzzy believe distance measurement and a new support probability in the meaning of probability fusion, make full use of the advantages of the scope's certainty of membership function in the fuzzy set theory. Firstly, define the fuzzy support probability of the uncertain information. Then, obtain the credibility of the information measured by each sensor by using the membership function, quantization process the degree of trust between the data measured by sensors, measure the comprehensive degree of trust of the data measured by sensors though the trust matrix[6], use the above results to structure the mass function in the theory of evidence. Finally, using the D-S standards on multi-sensor information fusion, so that the fusion result more accurate and credible.

II. BACKGROUND

A Related Parameter Definition on the Support Degree of the Information

Hypothesis in a region spread n sensor nodes, the data measured from node I and node j are x_i and x_j , they are obey Gaussian distribution, and with their Gaussian curve as the characteristic function of the sensor, marked as $p_i(x)$, $p_j(x)$. The once observed value of X_i, X_j are x_i, x_j. Reflect the Size of deviation between x_i and x_j with the believe distance measurement, suppose as follow

Manuscript received May. 25, 2012;
Project number: JUDCF11003, 2009AA05Z203, XZZ11_0465,
* Corresponding author.

doi:10.4304/jsw.8.5.1157-1161

$$d_{ij} = 2\int_{x_i}^{x_j} p_i\left(x\middle|x_i\right)dx \quad (1)$$

$$d_{ji} = 2\int_{x_j}^{x_i} p_j\left(x\middle|x_j\right)dx \quad (2)$$

Where:

$$p_i\left(x\middle|x_i\right) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left\{-\frac{1}{2}\left(\frac{x-x_i}{\sigma_i}\right)^2\right\} \quad (3)$$

$$p_j\left(x\middle|x_j\right) = \frac{1}{\sqrt{2\pi}\sigma_j}\exp\left\{-\frac{1}{2}\left(\frac{x-x_j}{\sigma_j}\right)^2\right\} \quad (4)$$

d_{ij} express the believe distance measurement between the sensor I and the sensor j, obtain (5)with error function $erf(\theta)$ [7,8]\

$$d_{ij} = erf\left(\frac{\left|x_j - x_i\right|}{\sqrt{2}\sigma_i}\right), d_{ji} = erf\left(\frac{\left|x_i - x_j\right|}{\sqrt{2}\sigma_j}\right) \quad (5)$$

Suppose r_{ij} express the degree of x_i supported by x_j, so r_{ij} should be the function of believe distance measurement, such as (6), and with bigger of d_{ij}, smaller of r_{ij}. According to the general rule of people know something: for the difference of very big or very tiny are easy to distinguish, but for some case in between are generally very difficult to distinguish. So, According to a practical problems', specific background can be set specific believe threshold value δ_1, δ_2, and $0 \le \delta_1 < \delta_2$.

$$r_{ij} = \begin{cases} 1, & d_{ij} \le \delta_1 \\ 0, & d_{ij} \ge \delta_2 \\ f\left(d_{ij}\right), & \delta_1 < d_{ij} < \delta_2 \end{cases} \quad (6)$$

$f(x)$ is a function of strictly monotone down and continuous, when $x \in [\delta_1, \delta_2]$, we can obtain $0 \le f(x) \le 1$, $f(\delta_1) = 1$, $f(\delta_2) = 0$. Define the expression of $f(x)$ is（7）,confirm $\delta_3 \in (\delta_1, \delta_2)$, and insure $f(\delta_3) = 0.5$ is right, so the function can be defined:

$$f(x) = \begin{cases} 0.5 - \dfrac{x-\delta_3}{2(\delta_3 - \delta_1)}, & x \in (\delta_1, \delta_3) \\ 0.5 - \dfrac{x-\delta_3}{2(\delta_2 - \delta_3)}, & x \in (\delta_3, \delta_2) \end{cases} \quad (7)$$

B The Support Degree Matrix of Information

Exist mutual support matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix},$$

When there have n sensors measure the same parameter, where $r_{ii} = 1, i = 1, 2, \cdots, n$.

The real support degree of x_i can't decided by single r_{ij}, should reflected by the integrated support degree of $r_{i1}, r_{i2}, \cdots, r_{in}$. If the integrated support degree of x_i is r_i, then the authenticity of x_i is higher with the bigger of r_i, but r_i should meet the following two conditions:

$$0 \le r_i \le 1 \text{ and } \sum_{i=1}^{n} r_i = 1;$$

r_i should comprehensive about the full information of every sub-system $r_{i1}, r_{i2}, \cdots, r_{in}$, in the support degree system of x_i, meaning exist a group of negative $\alpha_1, \alpha_2, \cdots, \alpha_n$, make

$$r_i = \alpha_1 r_{i1} + \alpha_2 r_{i2} + \cdots + \alpha_n r_{in} \quad i = 1, 2, \cdots, n \quad (8)$$

In matrix form, it is

$$r = R\alpha \quad (9)$$

Where: $r = [r_1, r_2, \cdots, r_n]^T$, $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_n]^T$. Based of $r_{ij \ge 0}$ and the nature of the nonnegative matrices, we can know existing a biggest modulus eigenvalue λ, which makes $\lambda\alpha = R\alpha$, and $\alpha_i > 0$. At this point, $\lambda\alpha$ can be treated as the measurement of the integrated support degree of sensor nodes, meaning $r = \lambda\alpha$. By the above conditions, can obtain $r_i = \alpha_i / (\alpha_1 + \alpha_2 + \cdots + \alpha_n)$, where r_i is the degree of support by other sensors of the information measured by sensor i.

III. Proposed Algorithm For Information Fusion

A. Basic Connotation of D-S

Evidence theory is based on the "recognition framework Θ ", it defines a set function $m: 2^\Theta \to [0,1]$,meet $\sum_{A \subset \Theta} m(A) = 1$ and $m(\Phi) = 0$, m is called basic reliability distribution based on , recognition framework Θ ,if A belong to recognition frame Θ ,it is called the probability distribution function, that is function *mass* ,which reflects the degree of evidence support to proposition, according to the theory of evidence define probability function:

$$Bel(A) = \sum_{B \subset A} m(B) \quad (10)$$

Only in the probability function to describe the trust of a proposition is not enough, must introduce a value that doubted the degree of A, that is:

$$Dou(A) = Bel(\bar{A}) \quad (11)$$

$$Pl(A) = 1 - Bel(\bar{A}) \quad (12)$$

Dou is the doubted function of A ,says the doubtful degree of , Pl is called plausibility function, says the plausibility degree of A. $\left[Bel(A), Pl(A)\right]$ says the uncertain interval of the evidence; $\left[0, Bel(A)\right]$ says the completely confidence interval of the proposition A ;says really don't doubt interval of the proposition A. Literature [8] have proved, among $m(A)$, $Bel(A)$, and $Pl(A)$ is mutual certain, therefore, this article only use $m(A)$ to determine the credibility degree of proposition A.

B. *Fuzzy Membership and Probability Distribution Values*

Generally, there have m sensors in n models of domain of discourse, so for any recognizable object, each mode's membership [9] of m sensors are given respectively by themselves, as follow:

$$\begin{bmatrix} \mu_{1.A_1}(x) & \mu_{1.A_2}(x) & \cdots & \mu_{1.A_n}(x) \\ \mu_{2.A_1}(x) & \mu_{2.A_2}(x) & \cdots & \mu_{2.A_n}(x) \\ \vdots & \vdots & & \vdots \\ \mu_{m.A_1}(x) & \mu_{m.A_2}(x) & \cdots & \mu_{m.A_n}(x) \end{bmatrix}_{m\times n}$$

The basic probability distribution function of each model belonged to object x is:

$$m_i(A_j) = \frac{r_i \mu_{i.A_j}(x)}{\sum_{j=1}^{n} r_i \mu_{i.A_j}(x)} \tag{13}$$

Where: $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$. $m_i(A_j)$ says the basic probability distribution values of model A_j belonged sensor i.

C. *Synthesis Rules oF D-S*

The combination rule of D-S theory is called orthogonal and rules, with $\oplus$ says, suppose m_1 and m_2 are two independent basic probability distribution values about 2^{Θ}, Bel_1 and Bel_2 are two trust function of the same recognition framework, and m_1 and m_2 are the corresponding basic probability assignment, the 焦元 of Bel_1 is $A_1, A_2, \cdots, A_J$, the 焦元 of $B_1, B_2, \cdots, B_K$, obviously:

$$K = \sum_{A_{ii} \cap B_{jj} = \Phi} m_1(A_i) m_2(B_j) < 1 \tag{14}$$

To express the utterance:

$$m(C) = \begin{cases} 0 & C = \Phi \\ \dfrac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - K} & \forall C \subset \Theta \ C \neq \Phi \end{cases} \tag{15}$$

If $K = 1$, m_1 and m_2 can be considered incompatible, so can't combine basic probability assignment; otherwise, $m(A)$ can decide a basic probability distribution value, which called comprehensive probability distribution value. According to more evidence combinations, can fuse each two evidence using this combination rule.

D. *Description of Algorithm*

Through analysis of the fuzzy membership and evidence theory, the algorithm model as shown in figure 1 :

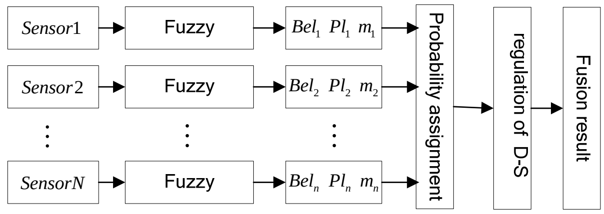

Figure 1 .Model of FS-DS fusion Algorithm

Algorithm steps can describe as follows:

(1)Use each sensor's local decision through converting the measured value of the each sensor to fuzzy value using membership functions;

(2) Calculate the support degree between each data. Create distance matrix firstly, then calculated mutual support degree matrix according to the distance matrix;

(3)Use the data that having measuring data calculate values fuzzy membership of the each sensor measured, and says the credibility of data information provided by each sensor with membership;

(4) Convert the degree of support and confidence to basic probability distribution function, and calculate the probability distribution value of each sensor;

(5) Fuse information using the D-S combination rule, and get the final result.

IV. Evaluation Of Proposed Algorithm

In this paper, the multi-sensor information fusion method based on fuzzy set and evidence theory is applied in the data vacancy of farmland soil moisture content[10], which achieved good results and enhanced the precision of data.

The soil is the source of nutrition for crop growth, and water is the most basic as well as the most essential nutrients. The accurately measured soil moisture content is of great significance to the growth of crop. And the soil moisture content generally refers to the absolute moisture content which means the contained moisture of 100G dry soil.

As shown in Table.1, the data of moisture content is multi-measured randomly by six different sensors from 20m below the surface of a test-bed whose area is 500.

TABLE 1.

Farmland soil moisture content

num	t_1	t_2	t_3	t_4	t_5	t_6
1	18.30	18.80	18.38	17.98	18.61	18.22
2	18.73	17.91	18.55	18.27	18.01	18.76
3	18.19	18.80	18.14	18.74	18.35	18.03
4	18.42	18.09	18.64	17.97	18.70	18.32
5	18.40	18.71	18.12	18.75	18.39	17.89
6	18.38	18.26	18.49	18.56	18.47	17.99

Because using different kinds of sensors to detect the information, so in order to get the credibility of each group of data, define the membership function as follow:

$$u(z) = \begin{cases} 1 - \dfrac{|z-\mu|}{2\sigma}, & |z-\mu| < 2s \\ 0 & |z-\mu| \geq 2\sigma \end{cases} \tag{16}$$

Where: z is the measured value; $\{\mu, \sigma\}$ is the mean value and deviation.

The credibility of measured data by each sensor can be calculated from expression (16), as shown table 2:

TABLE 2.

THE CREDIBILITY OF EACH GROUP OF DATA MEASURED BY EACH SENSOR

num	t_1	t_2	t_3	t_4	t_5	t_6
1	0.846	0.211	0.997	0.242	0.569	0.695
2	0.462	0.307	0.732	0.847	0.457	0.417
3	0.687	0.280	0.602	0.382	0.958	0.416
4	0.881	0.499	0.467	0.273	0.354	0.931
5	0.962	0.452	0.578	0.386	0.978	0.199
6	0.943	0.741	0.653	0.469	0.706	0.030

The believe distance matrix of each group measured data's support degree can be obtained from expression (5), as follow:

$$D=\begin{bmatrix} 0.0000 & 0.0240 & 0.0180 & 0.0749 & 0.0131 & 0.0978 \\ 0.0301 & 0.0000 & 0.0090 & 0.0450 & 0.0131 & 0.0560 \\ 0.0201 & 0.0080 & 0.0000 & 0.0550 & 0.0044 & 0.0700 \\ 0.0752 & 0.0359 & 0.0495 & 0.0000 & 0.0525 & 0.0070 \\ 0.0151 & 0.0120 & 0.0045 & 0.0600 & 0.0000 & 0.0769 \\ 0.0702 & 0.0319 & 0.0450 & 0.0050 & 0.0481 & 0.0000 \end{bmatrix}$$

Through calculating the expression (6) obtain each group data's support each other matrix, as follow (suppose δ_1=0, δ_2=0.06, δ_3=0.03):

$$R=\begin{bmatrix} 1.0000 & 0.6007 & 0.6997 & 0.2486 & 0.7812 & 0.0000 \\ 0.4983 & 1.0000 & 0.8499 & 0.2502 & 0.7812 & 0.0669 \\ 0.6655 & 0.8669 & 1.0000 & 0.0838 & 0.9271 & 0.0000 \\ 0.0000 & 0.4012 & 0.1747 & 1.0000 & 0.1253 & 0.8833 \\ 0.7491 & 0.8003 & 0.9249 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.4677 & 0.2497 & 0.9166 & 0.1981 & 1.0000 \end{bmatrix}$$

Through calculating can know the biggest model eigenvalue $\lambda = 3.3795$, and it's positive characteristic vector $\alpha = [0.4406 \;\; 0.4561 \;\; 0.4851 \;\; 0.2390 \;\; 0.4807 \;\; 0.2727]^T$, thus the comprehensive support degree of each group data $r = [0.1856 \;\; 0.1921 \;\; 0.2043 \;\; 0.1007 \;\; 0.2025 \;\; 0.1149]^T$,the degree of target attribute supported by each data(see table 3)can be obtained through consider the relation of each sensor , each sensor corresponding basic probability distribution values(see table 4) in each measured group data can be obtained by expression (13).

TABLE 3.

THE DEGREE OF TARGET ATTRIBUTE SUPPORTED BY EACH DATA

num	t_1	t_2	t_3	t_4	t_5	t_6
1	0.157	0.039	0.185	0.045	0.106	0.129
2	0.089	0.059	0.141	0.163	0.088	0.080
3	0.140	0.057	0.123	0.078	0.196	0.085
4	0.089	0.050	0.047	0.028	0.036	0.094
5	0.195	0.091	0.117	0.078	0.198	0.040
6	0.108	0.085	0.075	0.054	0.081	0.003

TABLE 4.

EACH GROUP DATA'S BASIC PROBABILITY DISTRIBUTION VALUE ABOUT EACH SENSOR

num	t_1	t_2	t_3	t_4	t_5	t_6
1	0.238	0.059	0.280	0.068	0.160	0.195
2	0.143	0.095	0.227	0.263	0.142	0.130
3	0.207	0.084	0.181	0.115	0.288	0.125
4	0.259	0.146	0.137	0.080	0.104	0.274
5	0.271	0.127	0.163	0.108	0.275	0.056
6	0.266	0.209	0.184	0.133	0.199	0.009

It is known that the degree of measured data supported by other groups from high to low in the order 3,5,2,1,6,4,through the comprehensive support degree. First to fuse the group of 3 and 5, fusion of the results:

$$m_3 \oplus m_5(t_1) = 0.287$$
$$m_3 \oplus m_5(t_2) = 0.055$$
$$m_3 \oplus m_5(t_3) = 0.151$$
$$m_3 \oplus m_5(t_4) = 0.064$$
$$m_3 \oplus m_5(t_5) = 0.407$$
$$m_3 \oplus m_5(t_6) = 0.036$$

Fuse the above results and the second data can obtain:

$$m_3 \oplus m_5 \oplus m_2(t_1) = 0.257$$
$$m_3 \oplus m_5 \oplus m_2(t_2) = 0.033$$
$$m_3 \oplus m_5 \oplus m_2(t_3) = 0.215$$
$$m_3 \oplus m_5 \oplus m_2(t_4) = 0.105$$
$$m_3 \oplus m_5 \oplus m_2(t_5) = 0.361$$
$$m_3 \oplus m_5 \oplus m_2(t_6) = 0.029$$

Fuse the above results and the 1, 6, 4, group in turn, finally can receive:

$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_1) = 0.595$$
$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_2) = 0.008$$
$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_3) = 0.215$$
$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_4) = 0.011$$
$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_5) = 0.169$$
$$m_3 \oplus m_5 \oplus m_2 \oplus m_1 \oplus m_6 \oplus m_4(t_6) = 0.002$$

Through the fusion results it was clear that the credibility of the sensors, That is the measured data of sensor t_1 had the highest accuracy, sensor t_3 taken second place. So should set up a threshold of basic probability distribution value to abandon the small data. Because the data having big deflection would impact the measured information, so the threshold could greatly abate the impact. In this paper, threshold δ=0.2 , that is give up the information measured by sensors whose probability distribution value less than 0.2.

Above analysis, the information measured by sensor t_1 and t_3 are credible information, combined with the comprehensive support degree of each group measured data could obtain the farmland soil moisture content value is 18.3775. For the purpose of data fusion is to get the real value of measured information, so compared with the ultimate value of soil moisture content in literature [11], the sample standard deviation obtained by data fusion method proposed in this paper greatly decreased to 0.0070, it is to appear more superior when compared to arithmetic average method or general trust function method, because their sample standard deviations are 0.0102 and 0.0094 respectively.

When multi-sensor measures the target information, environment and other various external factors may affect it, but if the received data of sensor is serious distortion then it can't recognize with the method proposed in literature [11], leading to finally fusion data is far from the real value. Because of the method without considering the relationship between sensors, and sensor in the calculation of the support degree too absolute and subjective, so leading to the error too large.

The FS-DS algorithm proposed in this paper fuzzy mutual support degree between different data avoid absolution of support each other between two data, so that would not change the mutual support relationship two between data when one part of the data had a small change, and ensure that information fusion have a better stability and anti-interference performance.

V. Conculusion

When multi-sensor perform the measurement towards a certain parameter, they will get different results because of the influence of the environmental factor and sensor itself. When data fusion, ensuring the reliability of measured data by multi-sensors is of great importance . The proposed FS-DS information fusion method in this paper is simple and easy to realize. This method could not only avoid the limitation of single sensor, but also reduce the influence of uncertainty error of sensors. What's more, this method can reflect the credibility of the measured information objectively and can alleviate the impact of larger deviation on data fusion results which can also avoid the absolute solution in the process of calculation .The real application results verify the validity of this method and enhance the precision of fusion results.

VI. Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (JUDCF11003), Jiangsu Innovation Program for Graduates (CXZZ11_0465) and Development Program of China (863) under Grant No 2009AA05Z203

References

[1] Yuan Xiaoguang, Yang Wanhai, Shi Lin. Multi-sensor Decision Fusion Design Based on Immune Strategy[J]. Journal of university of electronic science and technolgy of china. 2009,38(1):pp. 117-121.

[2] Xie Chun li, Guan Qiang. Research on Data Fusion Diagnosis System Based on Neural Network and D-S Evidence Theory[C]//Proc of IEEE Circuits and Systems International Conference on Testing and Diagnosis,Chengdu, CN, 2009: pp. 1-4.

[3] Guo Huidong, Zhang Xinhua, Song Yuan, Lu Qiangqiang. Asynchronous Multisensor Data Fusion [J]. Journal of Electronics&Information Technology, 2006,28(9): pp.1546-1549.

[4] Wang Yaonan, Li Shutao. Multisensor information fusion and its appplication[J]. Control and decision, 2001,16(5):pp. 518-522.

[5] Hu Zhentao, Liu Xianxing. Method of multi-sensor data fusion based on relative distance [J]. Systems Engineering and Electronics, 2006,28(2): pp.196-198.

[6] Zhang Changfang, Yang Hongwen, Hu Wwidong. Many-to-one Matching decision based on correction of perturbation for matching measure matrix[J]. Signal Processing.2009, 25(12):pp. 1908-1911.

[7] Daniel Cobb.On the solution of linear differential equations with singular coeffici-ents[J]. Journal of Differential Equations, 1992,46(2): pp. 310-323.

[8] Bruce A Francis.Convergence in the boundary layer for singular perturbed equations [J]. Automatica. 1982,18(2): pp. 57-62.

[9] Han Feng, Yang Wanhai, Yuan Xiaoguang. Evidence theory information fusion method based on fuzzy set [J]. Control and Decision, 2010,25(3): pp.449-452.

[10] Zhen Li, Wang Ning, Hong Tiansheng,et al. Design of wireless sensornetwork system based on in-field soil water content monitoring[J]. Transactions of the CSAE,2010,26(2):pp. 212-217.

[11] Jiao Zhuqing, Xiong Weili, Xu Baoguo. Multi-sensor data fusion method based on belief degree and its applications [J]. Journal of Southeast University, 2008,38(1): pp. 253-257.

Shunyuan Sun, born in1984, is a Ph.D. Candidate of Jiangnan University. His main research interests include Localization, channel estimation and equalization, detection, and cross-layer design and throughput analysis for wireless sensor networks